Université de Paris 8 & Laboratoire de Sciences Cognitives et Psycholinguistique,
EHESS-ENS-CNRS - www.lscp.net/persons/peperkamp

Statistical inferences and linguistic knowledge in early phonological acquisition

Sharon Peperkamp

Emmanuel Dupoux, Rozenn Le Calvez,
Jean-Pierre Nadal, Katrin Skoruppa

Workshop on Phonological Systems and Complex Adaptive Systems,
July 4-6 2005, Lyon

Early language acquisition: evidence for statistical inferences

- *segments*: distribution of tokens within the acoustic space (Maye, Werker & Gerken 2002)
  - exposure: monomodal or bimodal [ta]-[da] continuum
  - testing: discrimination of [ta]-[da]

- *phonotactics*: distribution of segments in onsets vs. codas (Chambers, Onishi & Fischer 2002; Saffran & Thiessen 2003)
  - exposure: CVC syllables with different sets of onset and coda consonants
  - testing: listening time for new syllables in which the consonant phonotactics are respected or not

- *word segmentation*: transitional probabilities (Saffran, Aslin & Newport 1996)
  - exposure: continuous speech stream consisting of 4 trisyllabic non-words (tupirobidakupadotigolabubidaku...)
  - testing: listening time for words (bidaku) and part-words (kupado)

Early language acquisition: evidence for linguistic inferences

- *segments*: generalization within a natural class (Maye & Weiss 2003)
  - exposure: monomodal or bimodal [ta]-[da] continuum
  - testing: discrimination of [ka]-[ga]

- *phonotactics*: better learning in case of natural classes than unnatural classes (Saffran & Thiessen 2003)
  - natural classes: /p,t,k/ in onsets, /b,d,g/ in codas
  - unnatural classes: /p,d,k/ in onsets, /b,t,g/ in codas

This talk

- Examine the respective roles of statistical and linguistic interferences for the acquisition of underlying representations

- Two complementary approaches
  - *modeling*: simulation on phonetically-transcribed speech
  - *experiments*: artificial language-learning paradigm

Acquisition of underlying representations

- establish *phoneme* inventory:

Spanish

	bilabial	labiodental	dental	alveolar	postalveolar	palatal	velar
stops	p b		t d				k g
fricatives		f v	θ	s			x
nasals	m		n			ɲ	
trills				r			
flaps			ɾ				
affricates					tʃ dʒ		
laterals				l		ʎ	

Acquisition of underlying representations

- establish *phoneme* inventory:
- on the basis of a *segment* inventory:

Spanish

	bilabial	labiodental	dental	alveolar	postalveolar	palatal	velar
stops	p b		t d				k g
fricatives	β	f v	θ ð	s z			x ɣ
nasals	m	ɱ	n			ɲ	ŋ
trills				r			
flaps			ɾ				
affricates					tʃ dʒ		
laterals				l		ʎ	

Acquisition of underlying representations

- establish *phoneme* inventory:
- on the basis of a *segment* inventory:

Spanish

	bilabial	labiodental	dental	alveolar	postalveolar	palatal	velar
stops	p b		t d				k g
fricatives	β	f v	θ ð	s z			x ɣ
nasals	m	ɱ	n			ɲ	ŋ
trills				r			
flaps			ɾ				
affricates					tʃ dʒ		
laterals				l		ʎ	

Acquisition of *segment* inventory

- age: between 6-12 months (Polka & Werker 1994; Werker & Tees 1984)
- method: prototype formation (Kuhl 1991; Kuhl *et al*. 1997; Maye, Werker & Gerken 2002)

Acquisition of *phoneme* inventory

- age: unknown
- method:
  - semantics

    [eldisko] ‘the disk’

    [miðisko] ‘my disk’
  - distributional analysis

    [ð]: intervocalically

    [d]: elsewhere
- objective: test the feasibility of the distributional mechanism
  - algorithm: look for complementary distributions of segments

A statistical algorithm

- Problems with basic algorithm
  - not robust to noice (production and/or perception errors)
  - fails to detect optional rules
- Solution: look for near-complementary distributions
  - for each segment, list the contexts in which it appears
  - for each pair of segments, compare the distributions of their contexts, by means of the Kullback-Leibler dissimilarity measure:

$$m_{KL}(s_1, s_2) = \sum_c \left(P(c|s_1) \quad \log\left(\frac{P(c|s_1)}{P(c|s_2)}\right) + P(c|s_2) \quad \log\left(\frac{P(c|s_2)}{P(c|s_1)}\right)\right)$$

s: segment c: context

A statistical algorithm

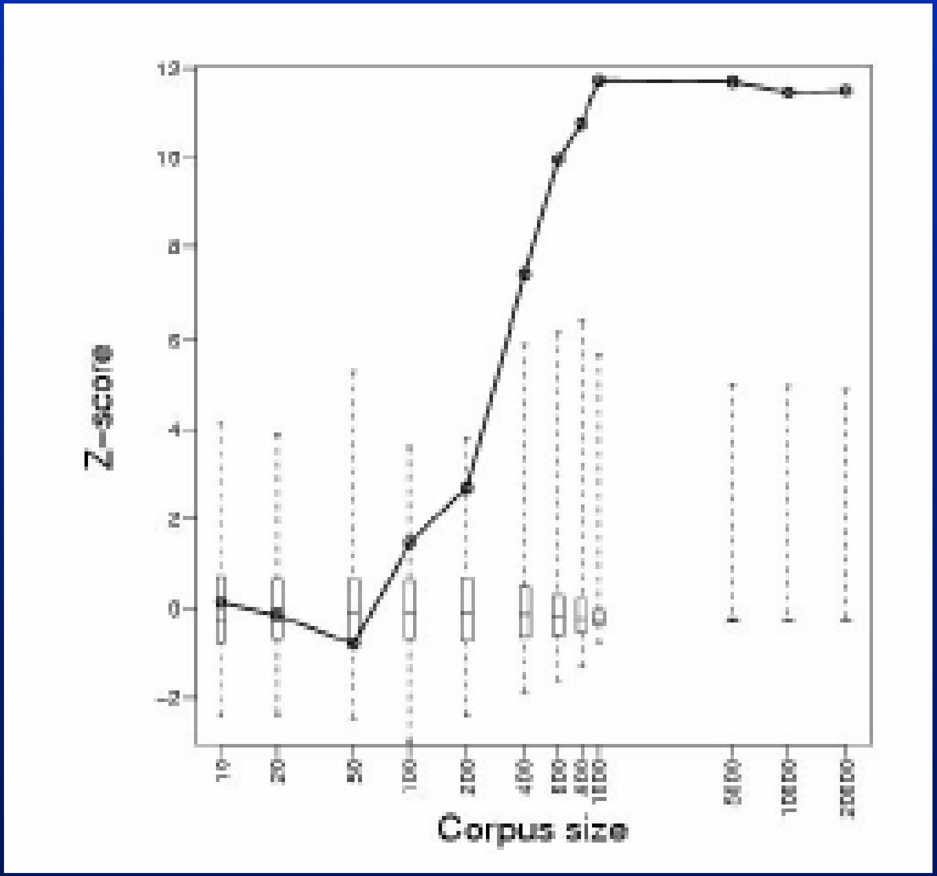
- For segment pairs with a KL number above some threshold, determine the default phone
  - the default segment is more frequent and appears in more contexts than the allophone
  - criterion of relative entropy:

$$s_d = \min_s \left[\sum_c P(c|s) \quad \log \frac{P(c|s)}{P(c)} \right]$$

s_d: default segment c: context

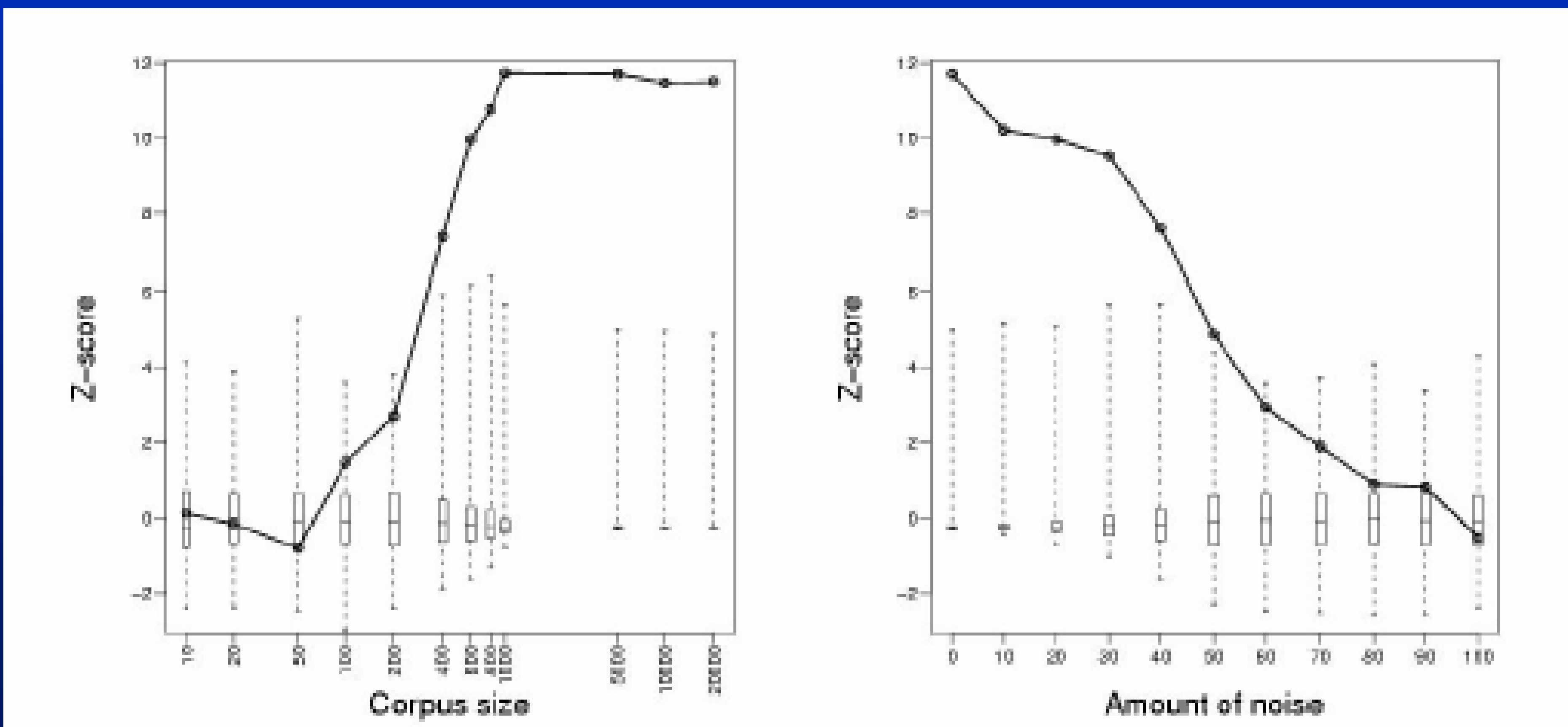
Validation on artificial corpora

- 46 phonemes with equal relative frequencies
- 1 phoneme has an allophone in 8 contexts
- utterances composed of random strings



Validation on artificial corpora

- 46 phonemes with equal relative frequencies
- 1 phoneme has an allophone in 8 contexts
- utterances composed of random strings



A statistical + linguistic algorithm

- Problem with statistical algorithm:
  - false alarms due to phonotactics (e.g. French: [�] only before vowels (*pluie*), [œ] only before consonants (*peur*))

- Solution: add a linguistically motivated filter
  - default phone and allophone are phonetically close
  - the context of a rule spreads a phonetic feature onto its targets

A statistical + linguistic algorithm

- Define each segment as a numerical vector encoding five articulatory properties (place, sonority, voicing, nasality, lip rounding)

- Criteria for detecting false alarms
  - there is a segment between default segment and the allophone:

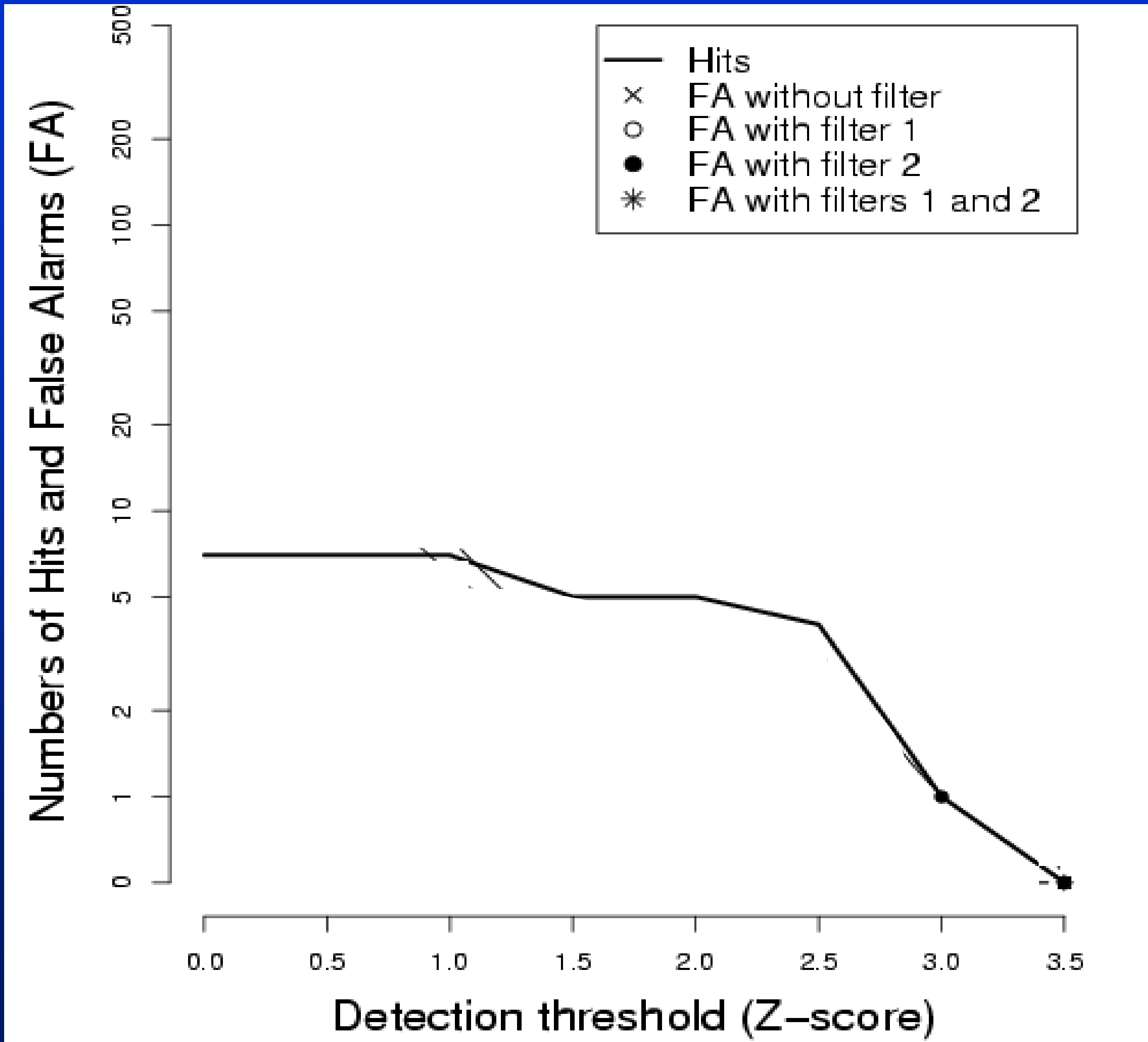
$$\exists s, \forall i \in \{1,..5\}, v_i(s_a) \le v_i(s) \le v_i(s_d)$$
$$or\ \forall i \in \{1,..5\}, v_i(s_d) \le v_i(s) \le v_i(s_a)$$

  - the allophone is more distant from its contexts then the default segment:

$$\exists i \in \{1,..5\}, \left| \sum_{s \in C[s_a]} \left(v_i(s_a) - v_i(s)\right) \right| > \left| \sum_{s \in C[s_a]} \left(v_i(s_d) - v_i(s)\right) \right|$$

Test on natural corpus

- CHILDES corpus
  - 42.000 short utterances of French parents to their children
  - transcribed phonemically
- Implementation of two allophonic rules:
  - palatalisation of /k/ and /g/ before /i,y,✵,☊,e,★,j,⯑/
  - devoicing of /r,l,m,n,✪,⬃,j/ before /p,t,k,f,s,▪/
- Corpus statistics:
  - Total number of segments: 35 default segments + (2+7) allophones = 44
  - Total number of segment pairs: 946



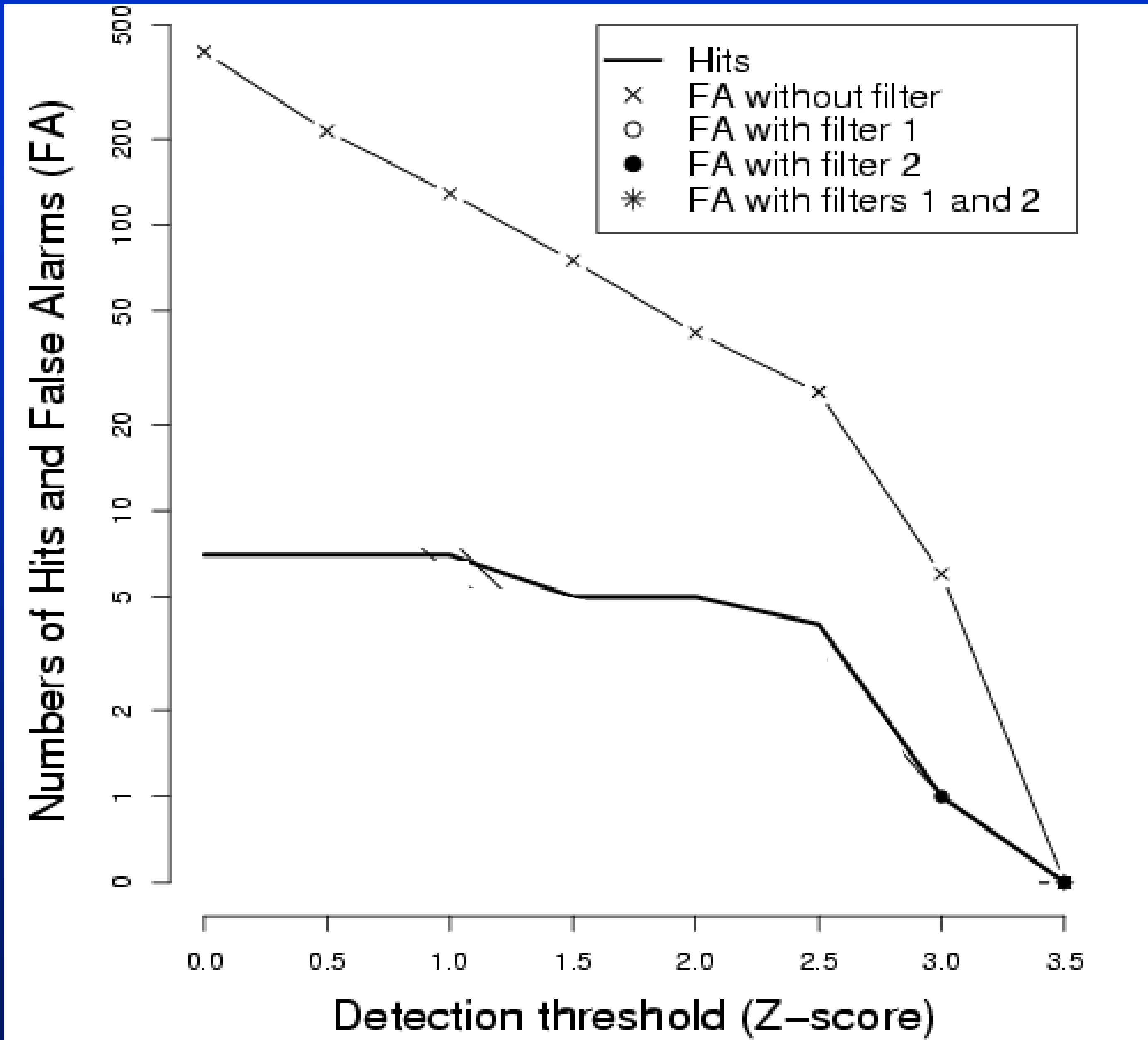
$0 \leq Z \leq 1$:

hits: 7

misses: 2

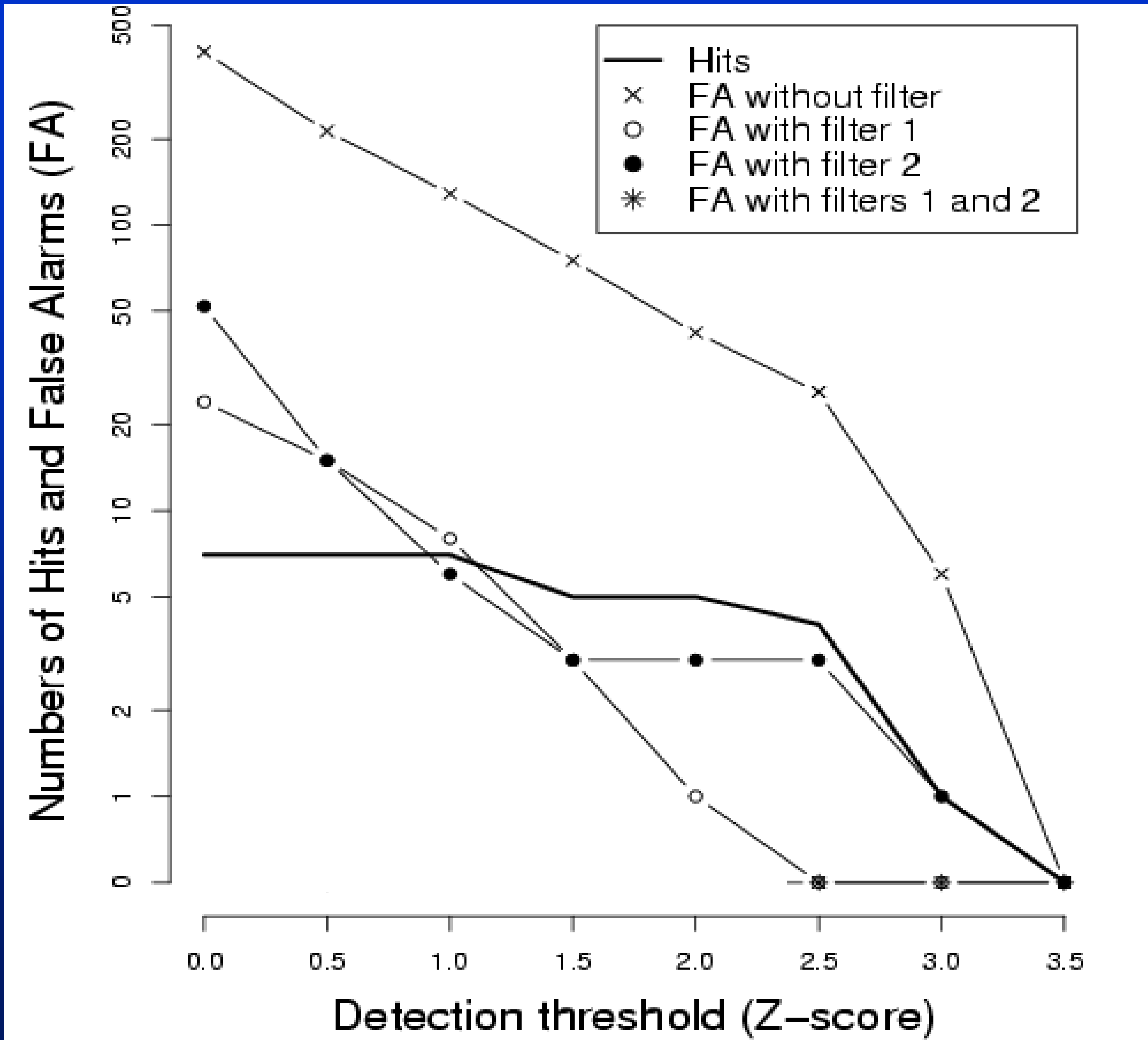
- /✪,⇙/, relative frequencies .02 and .01%

Peperkamp et al., submitted

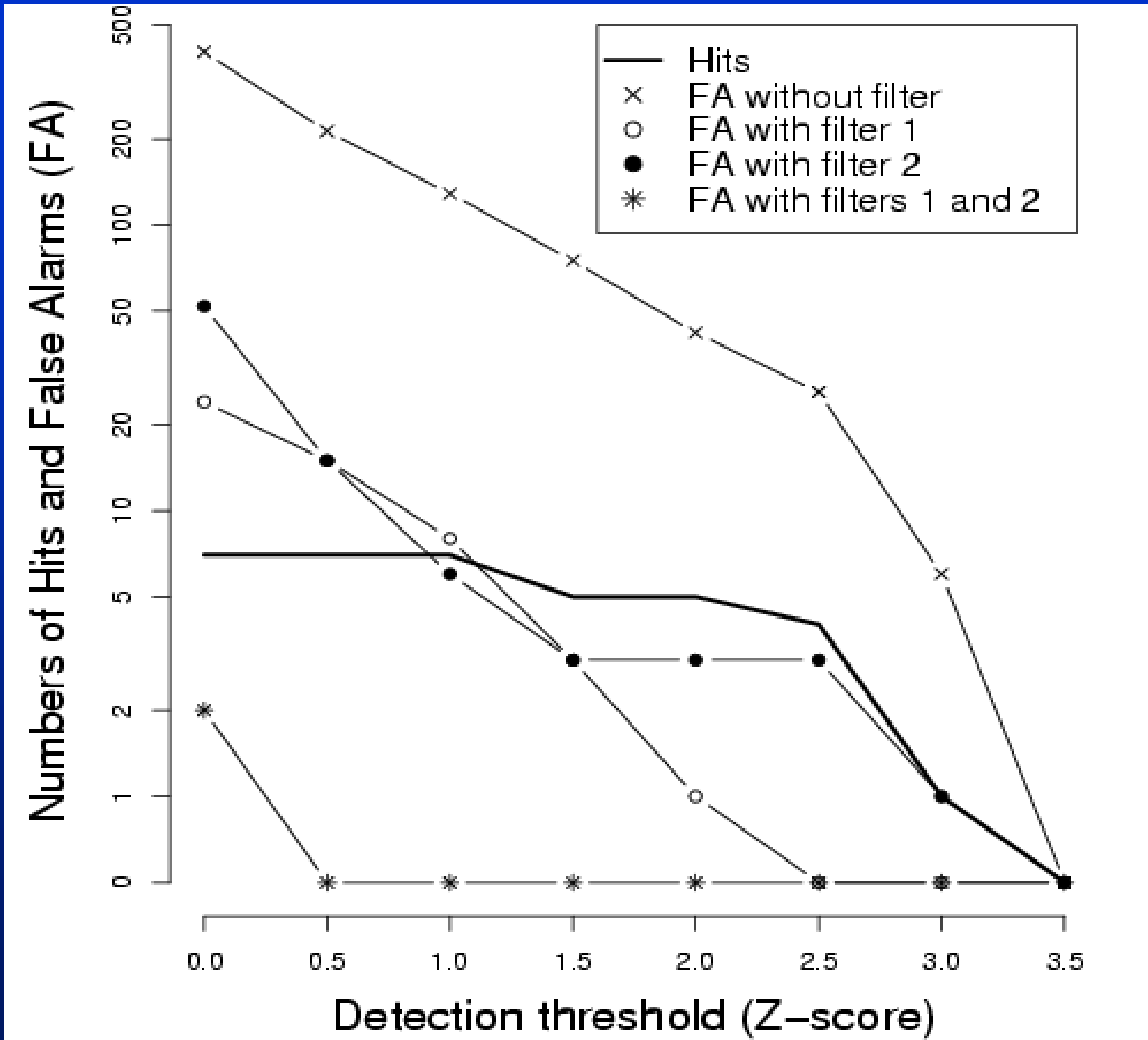


number of FA’s at Z=1: 129 (=13.6%)

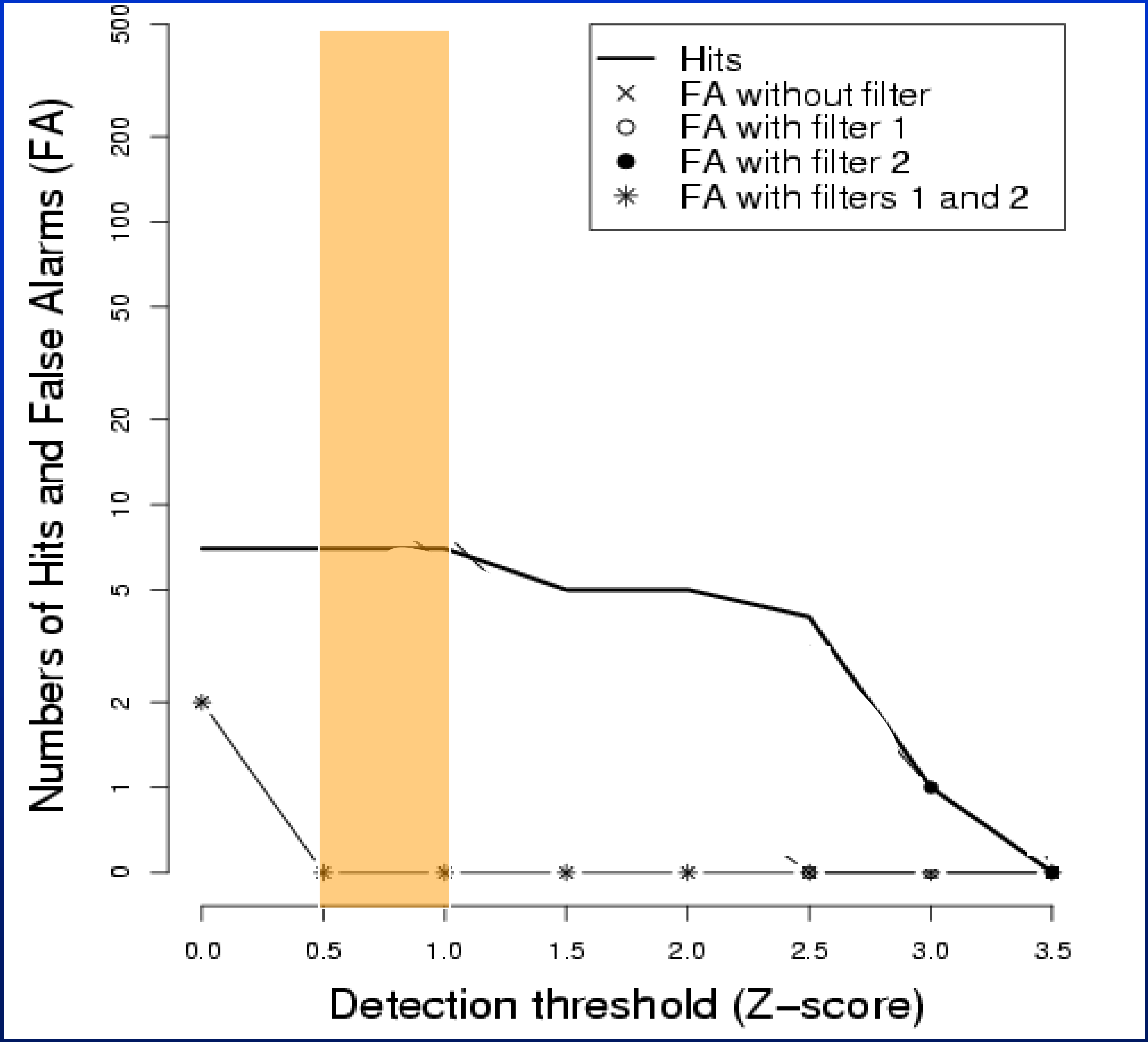
Peperkamp et al., submitted



Peperkamp et al., submitted



Peperkamp et al., submitted



$0.5 \leq Z \leq 1$:

- max. number of hits
- no FA's

Peperkamp et al., submitted

Summary

- Allophonic rules can be discovered in the absence of lexical knowledge on the basis of distributional information
- Linguistic knowledge concerning the nature of phonological rules is sufficient to discard false alarms
- Possible extensions
  - rule interaction
  - linguistic filter based on acoustic rather than articulatory distance (cf. Mielke 2005)
- Next step: test if infants are sensitive to complementary distributions and if it matters if the allophonic groupings are natural or not (work in progress with Jim Morgan)

Experiments

- Test if adults use linguistic knowledge when learning novel phonological rules

- Method: artificial language learning paradigm
  - natural versus unnatural allophonic rule

Experiment 1

- Two artificial languages:

	Language A	Language B
stops	/ p t k b d g /	/ p t k /
fricatives	/ f s ʃ /	/ f s ʃ v z ʒ /

Experiment 1

- Two artificial languages:

	Language A	Language B	French
stops	/ p t k b d g /	/ p t k /	/ p t k b d g /
fricatives	/ f s ʃ /	/ f s ʃ v z ʒ /	/ f s ʃ v z ʒ /

Experiment 1

- Two artificial languages:

	Language A	Language B	French
stops	/ p t k b d g /	/ p t k /	/ p t k b d g /
fricatives	/ f s ʃ /	/ f s ʃ v z ʒ /	/ f s ʃ v z ʒ /

- Two natural allophonic rules:

  Language A: intervocalic *fricative* voicing

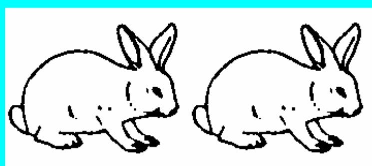
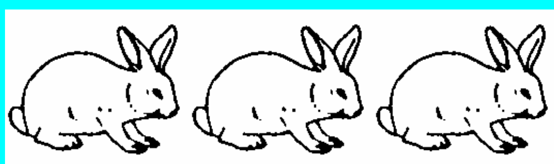
  Language B: intervocalic *stop* voicing

- Determinant + noun phrases:

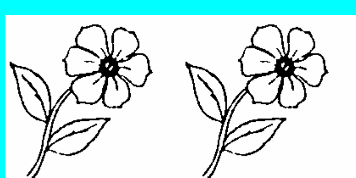
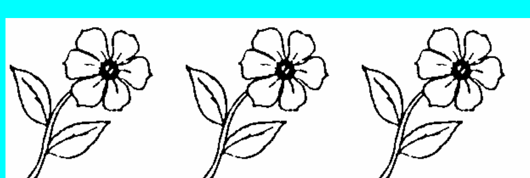
  nel ‘two’
  ra ‘three’

  nouns begin with a stop or fricative

Exposure: phrase-picture pairings

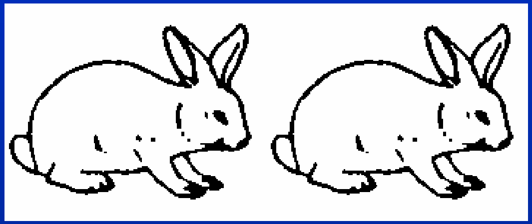
	Language A	Language B
	nel **p**emuʃ	nel **p**emuʃ
	ra **p**emuʃ	ra **b**emuʃ
	nel **f**oʒam	nel **f**oʒam
	ra **v**oʒam	ra **f**oʒam

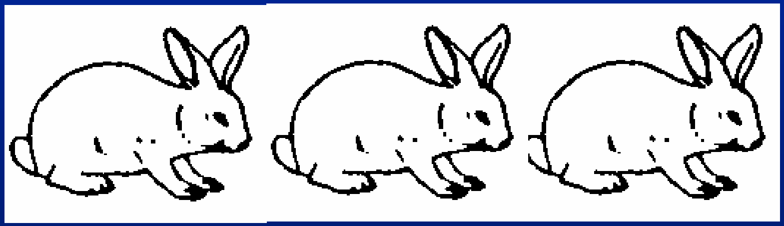
Exposure: phrase-picture pairings

	Language A	Language B
	nel **b**ovi	nel **p**ovi
	ra **b**ovi	ra **b**ovi
	nel **f**ulek	nel **v**ulek
	ra **v**ulek	ra **v**ulek

Test I: phrase production, known items

nel pemuʃ



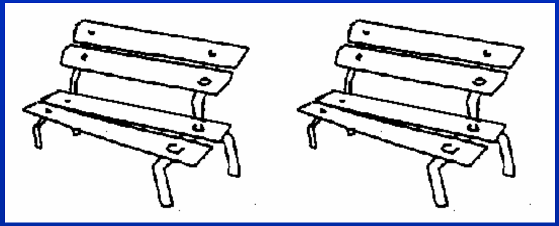


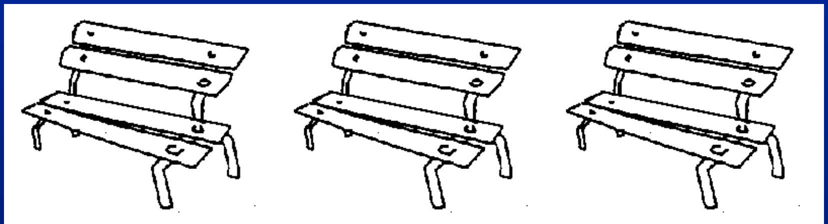
Language A: ra pemuʃ

Language B: ra bemuʃ

Test II: phrase production, new items

nel pura





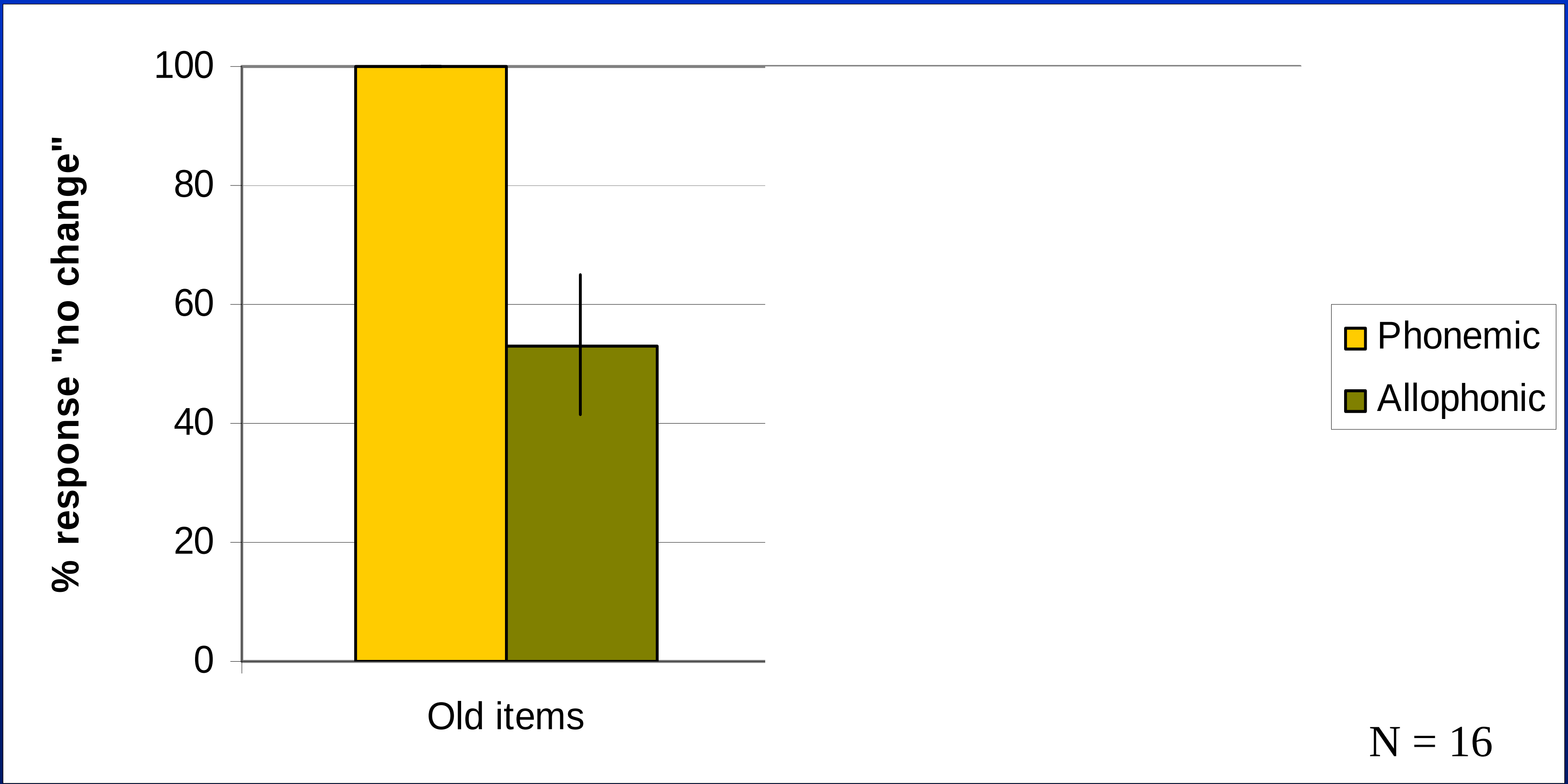
Language A: ra pura

Language B: ra bura

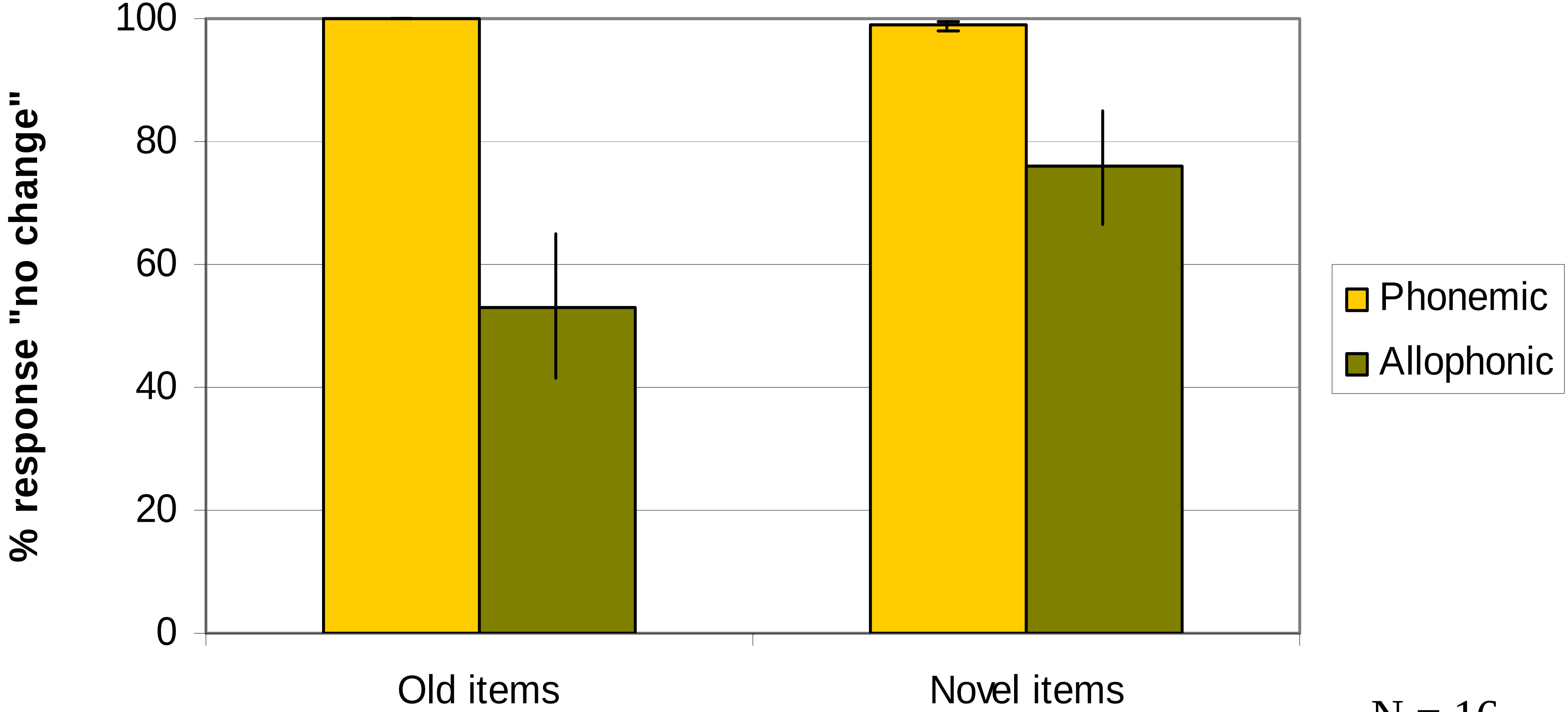
Experimental details

- Exposure phase (15 minutes):
  - – 8 lexical items
    - 4 stop-initial
    - 4 fricative-initial
  - – each item appears in 2 phrases (one with *nel* one with *ra*), repeated 16 times each
- Test phase:
  - – 8 old items
  - – 32 novel items

Results


100
80
60
40
20
0
% response "no change"
Old items
Phonemic
Allophonic
N = 16

Results


% response "no change"
100
80
60
40
20
0
Old items
Novel items
Phonemic
Allophonic

N = 16

Experiment 2

- Same segment inventories, different phoneme inventories:

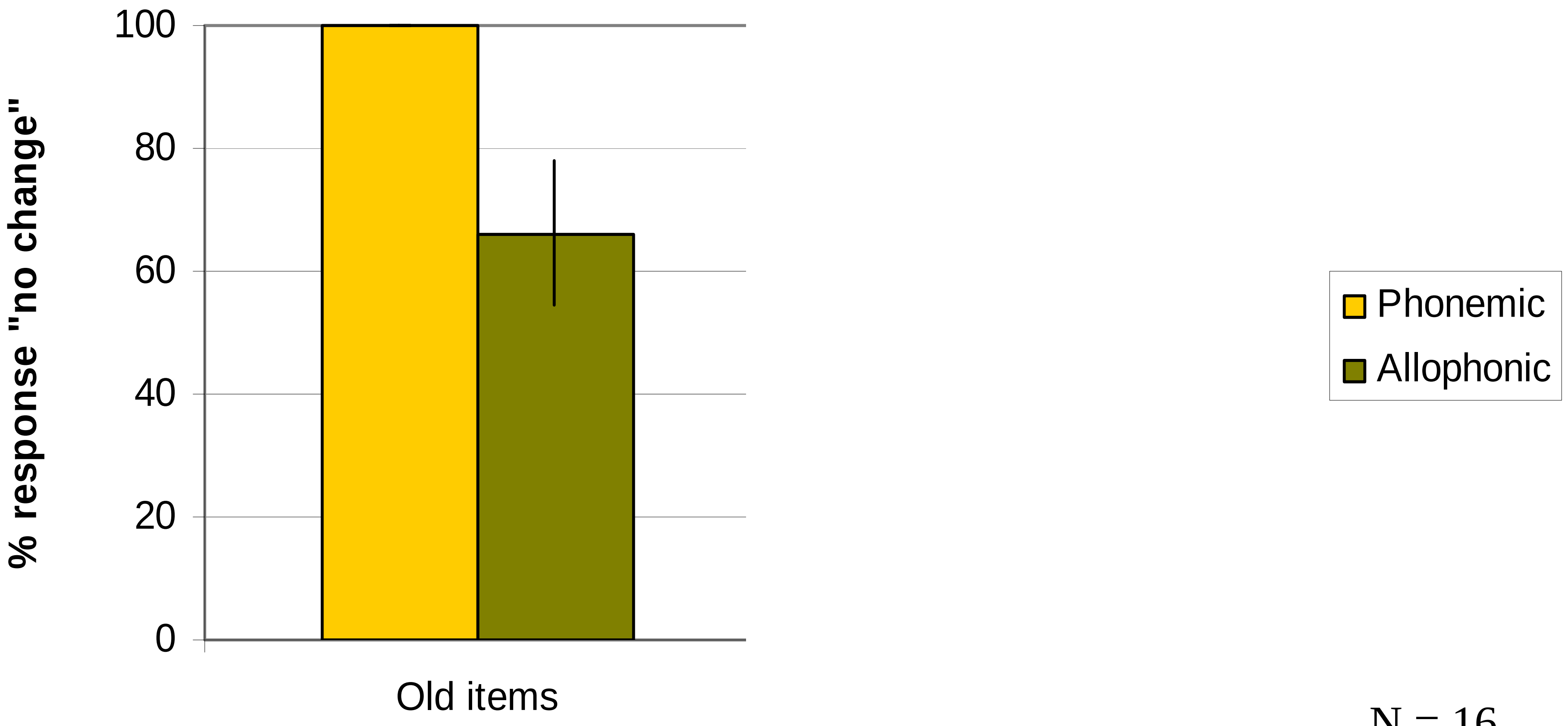
	Language A	Language B
stops	[p t k b d g]	[p t k b d g]
fricatives	[f s ʃ v z ʒ]	[f s ʃ v z ʒ]

- Two *unnatural* rules:

  Language A: /z/ → [t], /g/ → [f], /p/ → [ʒ] / V_V

  Language B: /v/ → [k], /ʃ/ → [b], /d/ → [s] / V_V

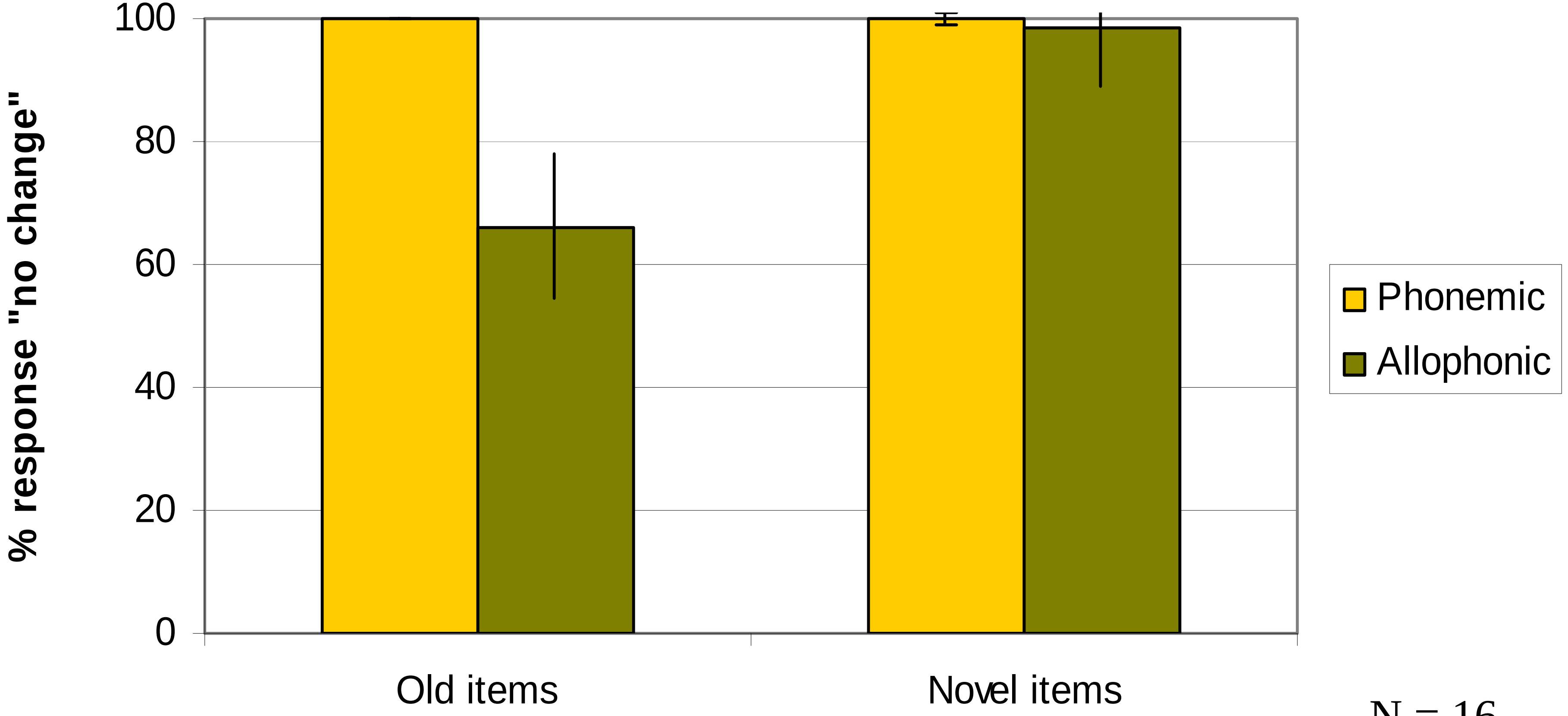
Results



100
80
60
40
20
0
% response "no change"
Old items
Phonemic
Allophonic

N = 16

Results


% response "no change"
100
80
60
40
20
0
Old items
Novel items
Phonemic
Allophonic

N = 16

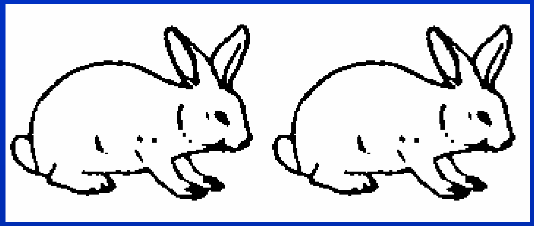
Summary

- Adults can learn the distinction between phonemic and allophonic contrasts within 15 minutes of exposure to an artificial language, but only if the allophonic groupings are phonetically natural

- Experiments 3 and 4:
  - as experiments 1 and 2, but with a perception rather than a production task

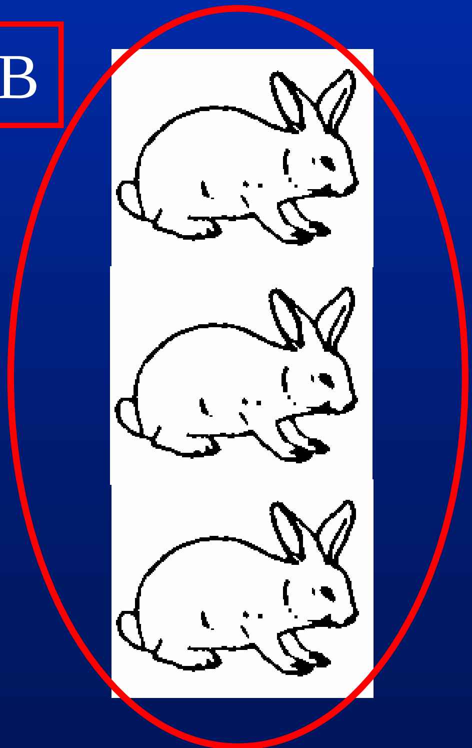
Test I: phrase-picture matching, known items

nel pemuʃ



ra bemuʃ

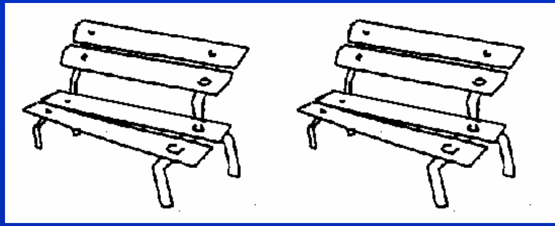
Language B



Language A

Test II: phrase-picture matching, new items

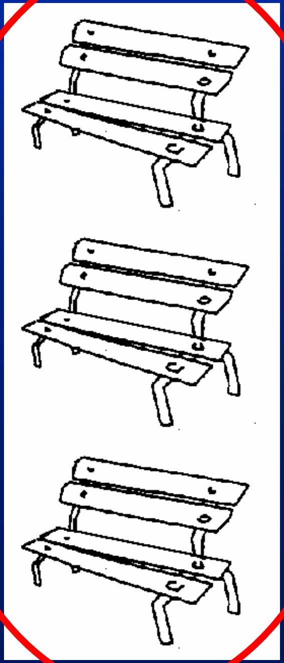
nel pura



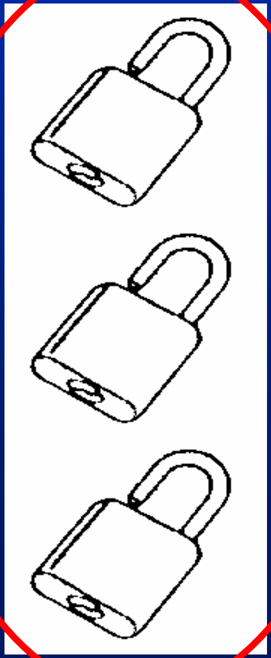
---

ra bura

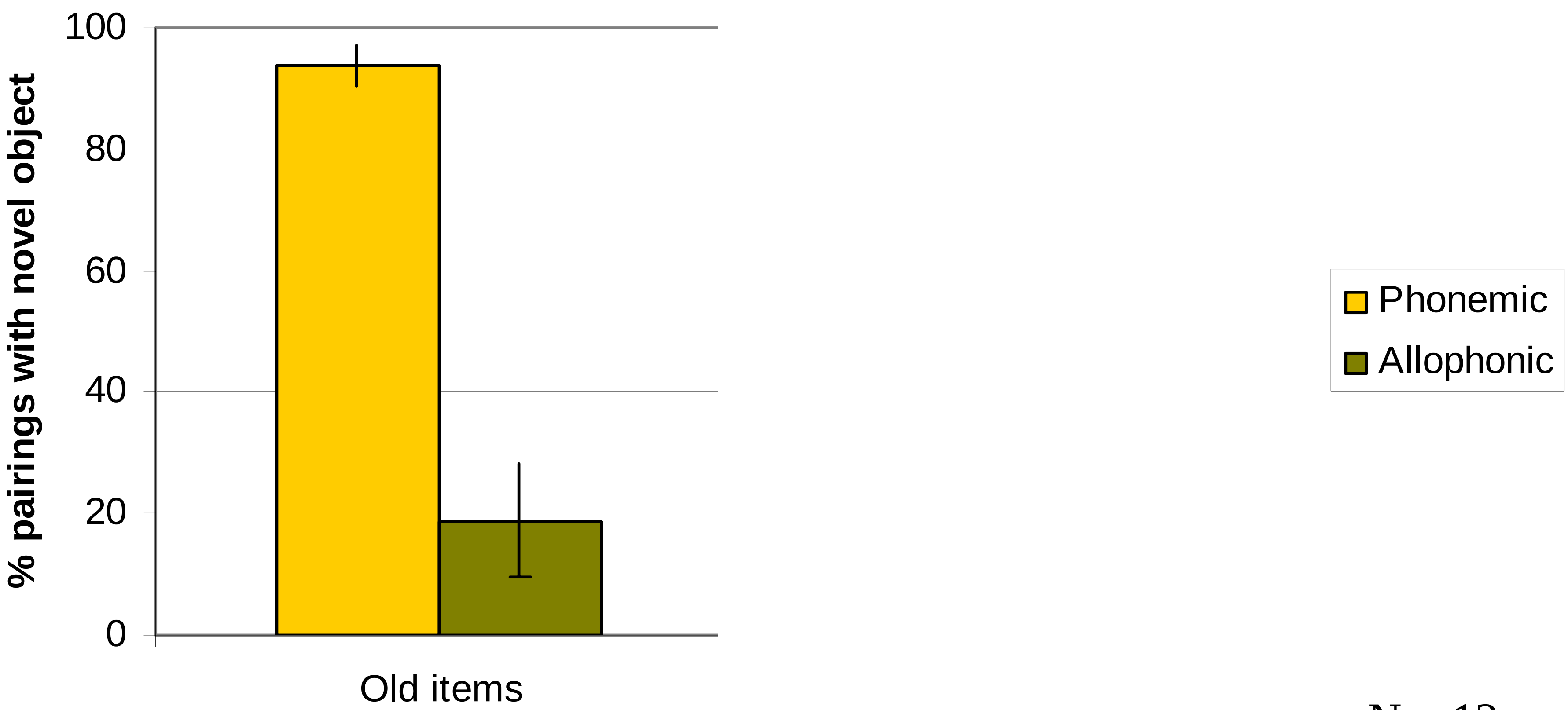
Language B



Language A



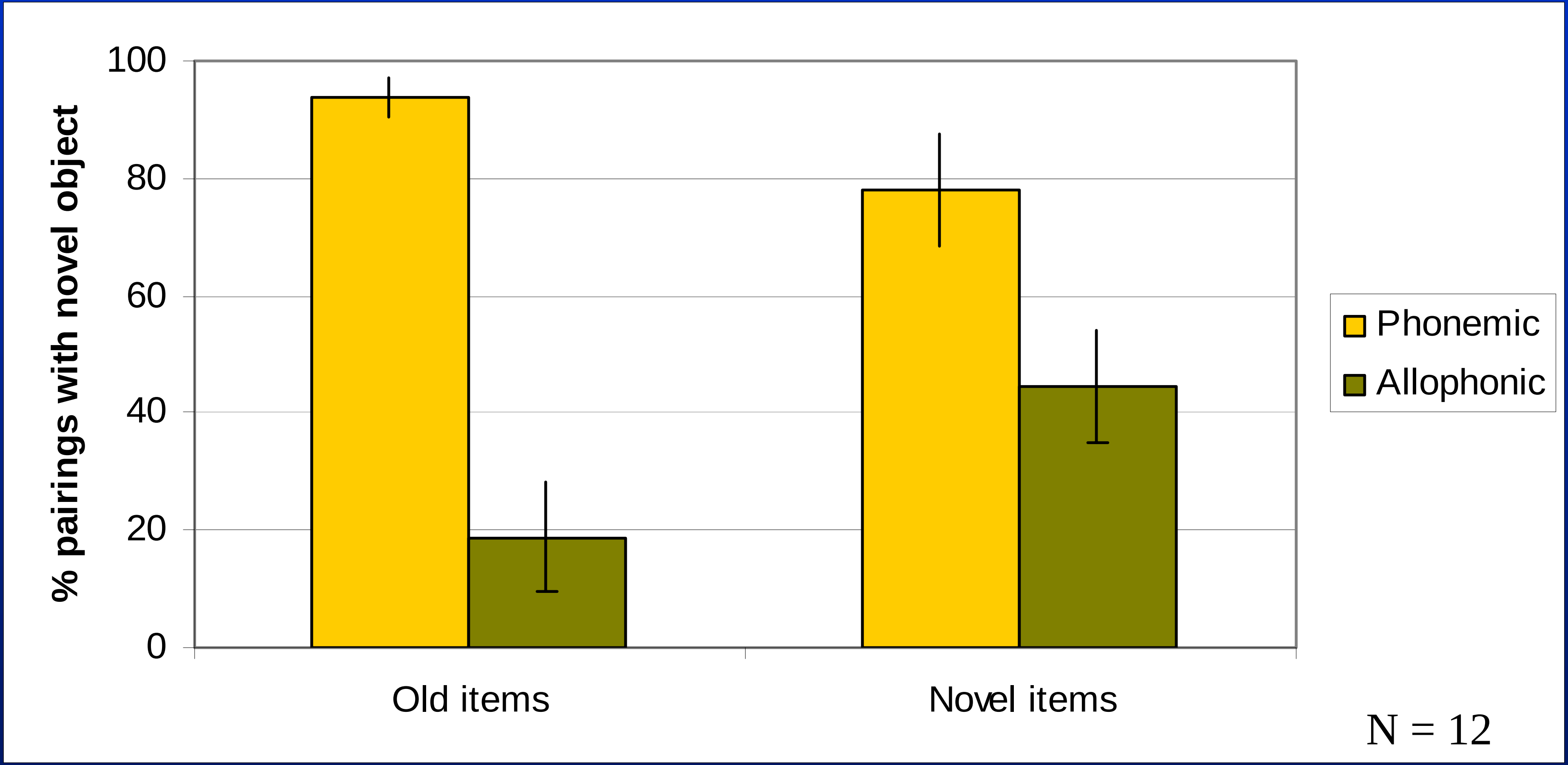
Results exp. 3: natural rule



N = 12

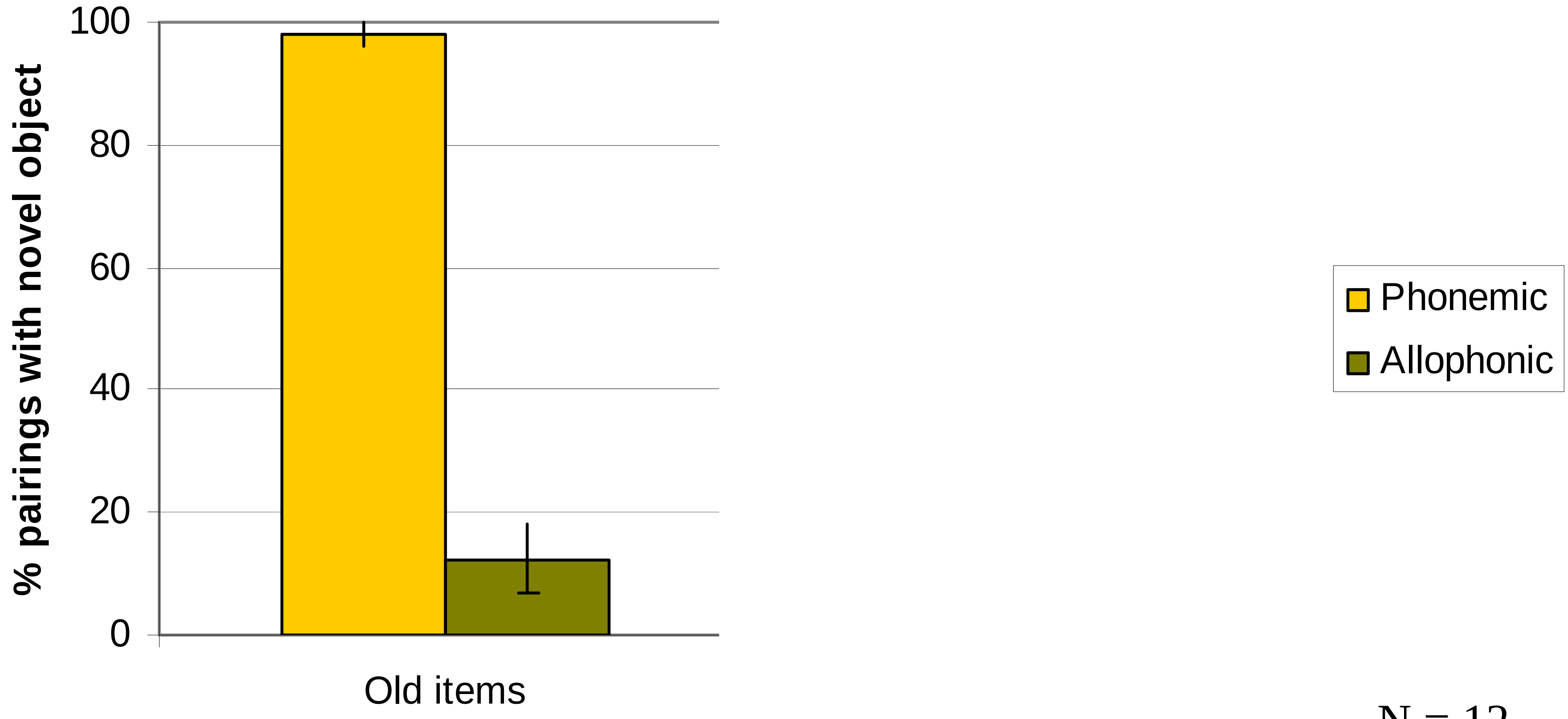
Peperkamp & Dupoux, to appear

Results exp. 3: natural rule



Peperkamp & Dupoux, to appear

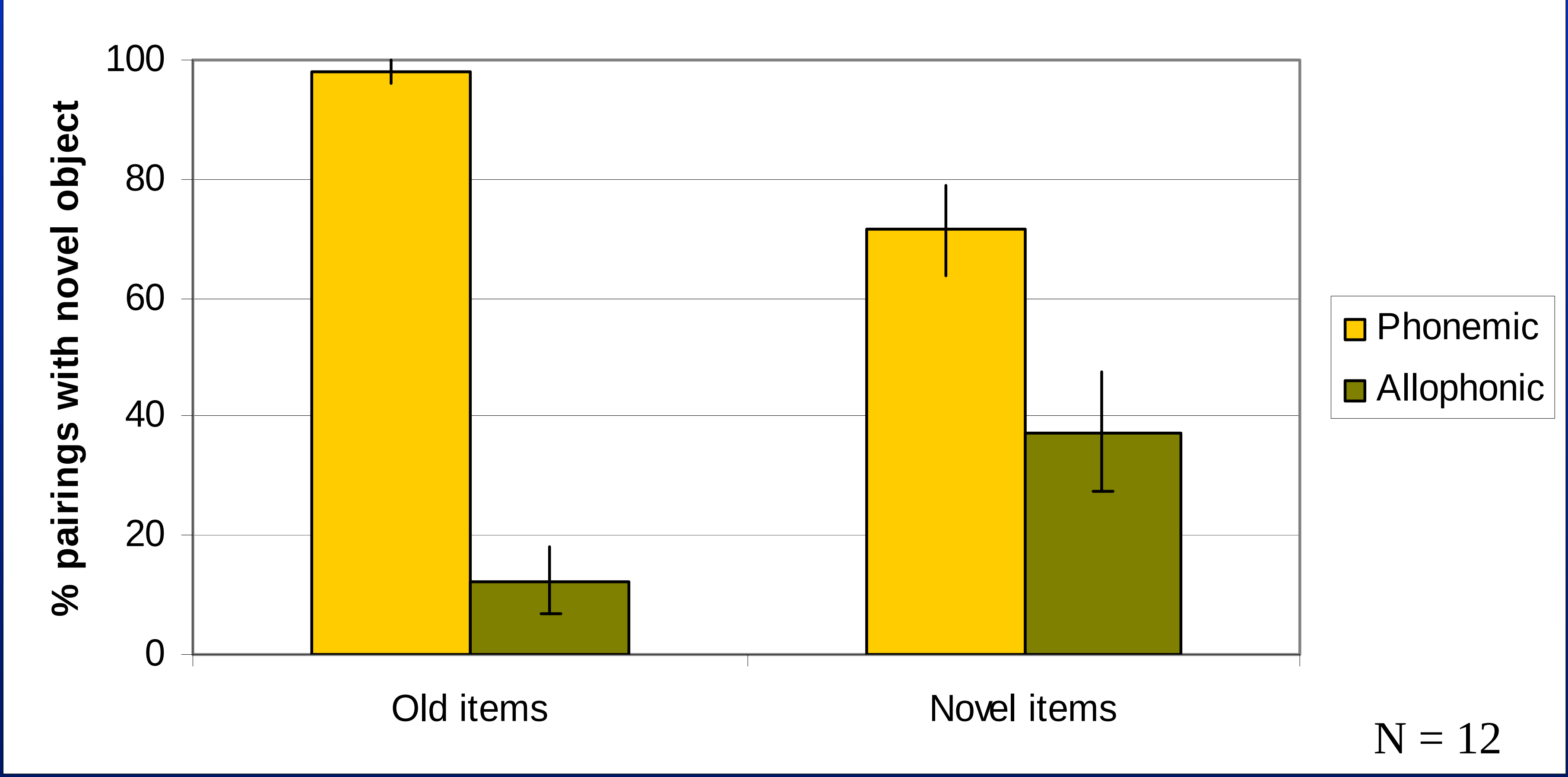
Results exp. 4: unnatural rule



N = 12

Peperkamp & Dupoux, to appear

Results exp. 4: unnatural rule



Peperkamp & Dupoux, to appear

Discussion

- Effect of phonetic naturalness with a production but not with a perception task

- Two possible explanations
  - task difference: free response in production, forced choice in perception

    → *make the perception task harder*
  - perception is not constrained by UG, but by a general algebraic learning system (Marcus *et al*, 1999)

    → *validate the present results with pre-school children and, ultimately, with infants*

Conclusion

- Both statistical and linguistic inferences seem necessary to acquire underlying representations

- Further research is necessary to
  - empirically demonstrate the presence of both types of inferences in infants
  - determine the precise nature of the linguistic inferences